JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

SIGN LANGUAGE RECOGNITION FOR DUMB AND DEAF

[1]Dr. C. Mahiba(Project Guide), Assistant Professor Department of Computer Science and Engineering, MVJ College of Engineering, India

[2345]Advika Shalini Naidu, Afnan S, Akshata Mudmal, Charitha Poonacha N, Department of Computer Science and Engineering, MVJ College of Engineering, India

Abstract : Communication plays a vital role in our day-to-day life. Humans communicate in several different ways out of which oral communication is the most prominent one. In the case of deaf and dumb people who cannot communicate orally, communication has been on a downside. For people with inability to communicate orally hands have been a major connection for communication, sing language is the method that is used to communicate through visual signs instead of spoken words. The sign language is difficult to understand by all. Across that globe 1.5 billion people that is 20% of the global population have hearing disability and around 5% of global population according to over 360 million people have lost their ability to speak as per the World Health Organization (WHO), the tools developed for communication for such people are much sophisticated equipment's such as smart gloves and hearing aids that come with higher cost and complexity. To overcome this and make communication easier and cheaper for the deaf-mute. This paper concentrates on a machine learning model that provides an interface for communicating through sign language that can convert sign language into text and speech. This machine learning model uses Convolutional Neural Network (CNN) that captures real time images through cameras in smartphones or web cameras in computer devices and provides the interpretation of the sign. The model aims at converting two major types of sign language American sign language (ASL) and Indian sign language (ISL). This paper deals with methods and technologies that could create a machine learning model that provides accurate and cost friendly communication for the impaired.

Key words: Graphical User Interface (GUI), Convolutional Neural Network (CNN), MATLAB, Indian Sign Language (ISL), American Sign Language (ASL), Gesture Recognition, Eigen value, Eigen Vector, Feature Extraction

I. INTRODUCTION

Communication is the method of exchanging information. Communication is of different types Verbal communication which uses words, sounds or speech to communicate, Non-Verbal communication doesn't use sound to communicate but uses physical signs or eye contact, Written communication is used to communicate using written or typed format, Listening is an important type of communication that aims at acquiring information than providing information and Visual communication uses visual elements to communicate. Verbal communication is the most important of all as it is done through spoken words, but in cases of people with disability of speaking or hearing, verbal communication is not possible. In such cases Sign languages are used to communicate.

Sign Language has been a vital tool for communication for the deaf-mute. Though sign language remains a prominent method for communication, it is still challenging for majority of the population to understand these sign languages. Two main Sign Languages we will be concentrating in this paper will be ISL - Indian Sign Language and ASL-American Sign Language. The Indian sign language comprises of around 300 signs and the American Sign Language consists of more than 10000 different signs including a-z and 0-9, with so many different signs it is difficult for a common man to understand these sign languages. To overcome this, sign language interpreters such a smart glove can be used but this requires equipment's that is costly to implement. This paper discusses a much cheaper and much more efficient method to interpret these sign language.

Digital Image Processing methods using Deep Neural Network can become an efficient method for interpreting sign languages. Convolutional Neural Network is currently the most famous and accurate image processing neural network. The digital images are converted into a large array of numbers with each pixel holding a certain numerical value. Image processing can be done for both grey scale images and color images. The image is read as a three-dimensional array in which the first two dimensions corresponds to the

height and width of the image and the last dimension of the array in a greyscale image is the value of the intensity of the light at that respective pixel and for color images the last dimension holds the RGB (Red, Green, Blue) value of an image. The images are analyzed based on the dot product of the image array and the filter array. The filters can be special filters that can find specific features in an image such as edges or patterns. The CNN typically has three important layers, convolutional layer, pooing layer and fully connected layer. The Convolutional layer is the core component of Convolutional Neural Network. This layer performs a dot product between two matrices that is the filter matrix also called as kernel, these filter matrixes are a set of learnable parameters and the image matrix which is an array of the numerical values of each pixel of the image that must be processed. The pooling layer performs pooling functions which reduces the spatial size of the dot product obtained in the convolutional layer. This reduces the number of learnable parameters and thus reduces the computational effort. Fully connected layer is the final layer of a Convolutional Neural Network that connects each input neuron to the output neuron.

Though convolutional Neural Network can be used to interpret sign language, it requires an interface that is handy and efficient. Android apps have been a great tool that can act as a cheap and efficient interface for deploying a machine learning model such as CNN. API's such as Flask and Flask Rest can be used to deploy the CNN model into websites and apps respectively thus making computers and mobile phones an easy and simple interpreter of Sign language in real time.

II. MOTIVATION

According to the World Health Organization 63 million people in India suffer from either complete or partial deafness. The speech disability rate in India according to the 2011 census in India reported that people with speech disability accounted for 7.6% and 6.1% of the total disability and in the census and survey, respectively. So, this project aims at providing free interpreters to aid disabled people and provide scope for interpreters at social places such as malls, airports, public buses etc. and aim at helping deaf-mute education.

III. PROBLEM STATEMENT

To implement an accurate and efficient deep learning neural network to recognize sign language in real time by capturing data through web cam or cameras available in computer devices and mobile phones respectively and also provide text to speech conversion of the recognized sign. Also, to provide a better, efficient and simple GUI for both websites and mobile apps.

IV. LITERATURE REVIEW

[1] Amit Dighe, this project builds a Convolutional Neural Network (CNN) model based on python using OpenCV which is a computer vision library. OpenCV is used to find the threshold value of each frame and return the contours. If contours are found in the ROI (Region of Interest) then the frame is saved as a hand sign and tested by the pretrained model. Once a sign is identified in the image it is classified as per the labels of the given training datasets.

[2] Mahesh Kumar N B, this project deals with sign language recognition of 26 Indian sign language signs using MATLAB. The model deals with preprocessing and hand segmentation, feature extraction, sign recognition and sign to text. It uses the Linear Discriminant analysis (LDA) algorithm for gesture recognition. Eigen values and Eigen vectors are extracted to recognize sign languages.

[3] Sharvani Amisha, this proposed model is used to detect sign language in Realtime taking inputs from webcam. The model is a tensor flow model that is used for sign language detection. This project makes use of python, Jupyter Notebook, OpenCV and TensorFlow object detection API and is built specially to detect Indian sign language.

[4] Shagun Katoch, the project aims at creating a model based on the technique of Bag of Visual Words (BOVW) to recognize Indian Sign Language alphabets (A-Z) and digits (0-9). The training data are labelled, and SURF (Speeded Up Robust) Features are extracted from the image to be tested and the Support Vector Machine (SVM) and Convolutional Neural Network (CNN) are used for Classifying the data. All these features can be accessed by a Graphical User Interface (GUI). Reverse recognition method for people to communicate back with the impaired is also implemented, in the reverse recognition the user inputs a specific letter or digit and the sign of the input entered is displayed on the screen.

[5] Ahmed J Abogarair, Image datasets are collected to train and test the Convolutional Neural Network. The data is labelled according to the signs in the image, the labelling of this data facilitates easy classification. All the images are resized to 128*128 pixels and are stored in RGB format. All the data that is used for training this model is captured using a webcam.

V. PROPOSED SYSTEM

A Convolutional Neural Network for Sign Language Recognition is built based on the following approach:

i. Data Acquisition

Image datasets are collected to train and test the Convolutional Neural Network. The data is labelled according to the signs in the image, the labelling of this data facilitates easy classification. All the images are resized to 128*128 pixels and are stored in RGB format. All the data that is used for training this model is captured using a webcam.

ii. Training

For training our Convolutional Neural Network (CNN), all the images are converted from RGB to greyscale thus reducing the training time and resource. Any Noise present in the image is removed by applying the Gaussian Blur. The images are resized to 128*128 pixels. After all the above-mentioned procedures the images if fed to the training model

In the training process the prediction layers predict as to which class the images belong to. The performance measurement factor for the training process is the cross entropy of the results from the prediction layer. This cross entropy is found using an in-built function in TensorFlow. When the cross entropy nears zero or equals to zero it means the image falls under that label.

iii. Image Frame Acquisition

For a sign to be recognized in an image the hands must be recognized, as the images are captured through the webcam of a computer or a smartphone camera, only the ROI (Region of Interest) that is the portion of the image that includes hands are cropped from the complete image. This is done using object identification, once the desired object is found only that part of the image is stored and processed.

iv. Image Pre-Processing

The image of the hands captured is converted from RGB to grayscale. Gaussian blur is applied over the image to remove any noise present in the image. The image is resized to 128*128 pixels. Thresholding is applied over the image making it convenient for feature extraction. All these can be achieved by the help of OpenCV library in python.

v. Feature Extraction

In this method the image data is differentiated into data that is important for the process if sign recognition, these data are also called features that facilitate gesture recognition, and the other data is the redundant data in the image that does not prove any usage in the process of sign language detection and is hence discarded. This step makes sure that computation occurs only on the required portion of the image data. This feature extraction is basically done through a method of contour-based shape recognition method. Using this method contours are recognized on images that provide the shape and other features in the image.

vi. Classification

Convolutional Neural Network (CNN) is used for classification of the image. The image is fed forward through the Neural Network as the input passes through the neural network it passes through the convolutional layer and the pooling layer where filters are passed in a sliding manner over the complete image to compare the features of an image with the required feature, that is required to classify the images according to their signs. The fully connected layer in Convolutional Neural Network connects each of the input neurons to the output neuron.

vii. Text to Speech

For much better and easier understanding of the sign the output is given in the form of text as well as speech. Several Text to Speech APIs are available in python, one such API is Google Text to Speech API that can be used to convert the text output from the Convolutional Neural Network into speech.

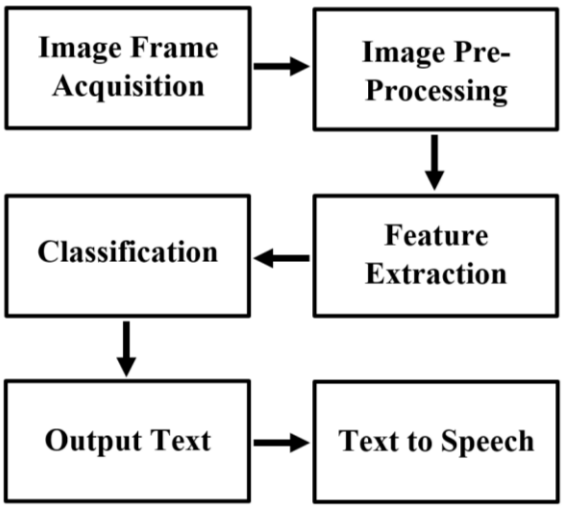

Fig. 1 – Flow Diagram of the Proposed Model

VI. Limitations and Challenges of the Proposed System

The challenges and limitations that generally rises with sign language recognition using deep learning network:

- Generating datasets: - The datasets required for training and validation is substantial and generating such huge datasets is a challenging task.
- Classification: - With sign language recognition several different filters are required to classify all most similar sign such as "P, R and D", "L, C and O", etc.
- Accuracy: - It is a challenging task to create a model with accuracy close to 99.9% and almost all models of sign language recognition present are around 95% to 97% accurate.
- Different sign languages: - Around the world there are more than different types of sign language each sign language requires a unique neural network model to recognize.
- Text to speech API: - Though hand gestures can be converted to text and speech. The problem is that around the world hundreds of different languages are used hence the text to speech API.

VII. Conclusion

In this paper we have discussed several methodologies to implement an interpreter for sign language. The proposed model is a much more accurate, efficient, and simple method compared to other methodologies proposed earlier. Though the proposed model covers all modern complications, it still has a lot of scope to be implemented in the future.

VIII. Acknowledgment

The authors would like to acknowledge and convey their warmest gratitude to the Head of the Department, Computer Science Engineering, MVJ College of Engineering and their project guide. They would also extend their heart felt thanks to their guardians and all those who supported them and helped them in this successfully completing this paper.

REFERENCES

[1] MR. Zubair Iqbal, Ritvik Tyagi, Ayush Kumr Singh, Aadarsh Upadhayay, Rishitha Chaudhary, "SIGN LANGUAGE TRANSLATIONS FOR DEAF AND DUMB", IJCRT, Volume 10 (May 2022))

[2] R Rumana, Reddygari Sandhya Rani, Mrs. R. Prema, "A Review Paper on Sign Language Recognition for The Deaf and Dumb", IJERT, Vol. 10 (October 2021)

[3] Manisha U. Kakde, Mahendra G. Nakrani, Amit M. Rawate, "A Review Paper on Sign Language Recognition System For Deaf And Dumb People using Image Processing", IJERT, Vol. 5 (March 2016)

[4] Nabah Inamdar, Zaid Inamdar, Sohel sheikh, Prof. Sunil A. Sushir, "A Survey Paper on Sign Language Recognition", IJRASET, Volume 10 (April 2022)

[5] Shruty M. Tomar, Dr. Narendra M. Patel, Dr. Darshak G. Thakore, "A Survey on Sign Language Recognition Systems", IJCRT, Volume 9 (March 2021)

[6] Satwik Ram Kodandaram, N Pavan Kumar, Sunil G L, "Sign Language Recognition", Turkish Journal of Computer and Mathematics Education, Vol.12 No.14(2021)

[7] Shagun Katoch, Varsha Singh, Uma Shanker Tiwary, "Indian Sign Language recognition system using SURF with SVM and CNN, ELSEIVER, (14 April 2022)

[8] Vijay More, Sanket Sangamnerkar, Vaibhav Thakare, Dnyaneshwari Mane, Rahul Dolas, "Sign Language Recognition Using Image Processing", (February 2018)

[9] A. Gayathri, Dr.A. Sasi Kumar, "SIGN LANGUAGE RECOGNITION FOR DEAF AND DUMB PEOPLE USING ANDROID ENVIRONMENT", ISSN, Volume-4 (2017)

[10] Ahmed Siltan, Walled Makram, Mohammed Kayed, Abdelmaged Amin Ali, "Sign language identification and recognition: A comparative study", DE GRUYTER, (2022)

[11] Prof. Radha S. Shirbhate, Mr. Vedant D. Shinde, Ms. Sanam A. Metkari, Ms. Pooja U. Borkar, Ms. Mayuri A. Khandge, "Sign Language Recognition Using Machine Learning Algorithm", IRJET, Volume: 07 (Mar 2020)

[12] Mahesh Kumar N B, "Conversion of Sign Language into Text", International Journel of Applied Engineering Research ISSN, Volume 13 (2018)

[13] Amit Dighe, Suraj Adsul, Saurabh Wankhede, Sonam Borhade, "Sign Language Recognition Application using Python and OpenCV", IJARSCT, Volume 2 (May 2022)

[14] Sharvani Srivastava, Amisha Gangwar, Richa Mishra, Sudhakar Singh, "Sign Language Recognition System using TensorFlow Object Detection API", International Conference on Advanced Network Technologies and Intelligent Computing (ANTIC-2021)